

# Double Dummy Hand Evaluator

Zhenyang Zhang  
Advisor: David Aldous

## 1. Introduction

The goal of this study is to analyze methods that already exist and come up with a new way of thinking about trick-taking values of different hands. Rather than making any conclusions on hand evaluation, I would like to be more cautious, and consider my model a “baby” version of a double dummy hand evaluator. I hope that this will bring up more thoughts about hand evaluation modeling in the future.

I am using double dummy data from Matt Ginsberg’s double dummy library. One may argue that double dummy tricks are not “realistic” since sometimes results are not achievable in the real world plays. For example, double dummy tricks will always run a suit missing Q, while in reality players may guess the Q on the wrong side and lose one trick as a result.

However, there are still a lot of advantages of using it. First, there are a lot of hands in Ginsberg’s library, 717102 hands in total. And in some sense double dummy tricks are more “pure”. We do not have to take into account if players made a mistake or used some at-table information (like hand gestures, facial expressions, etc.) during the play so that they won more or fewer tricks than expected. Also, expert players have inclination to bid games, which means that minor suit contracts will be less played and the result will be somehow biased, while double dummy results will not be affected.

## 2. Observations on existing methods

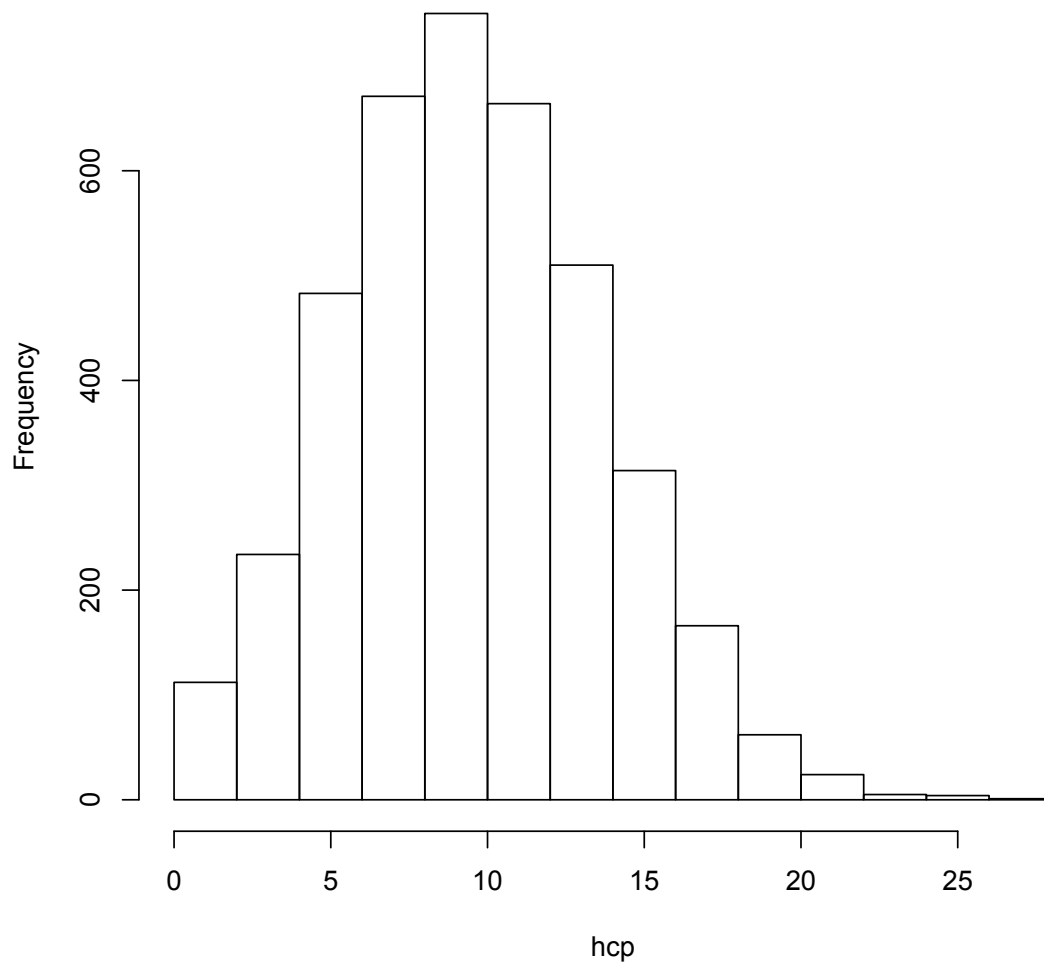
Before we actually step into our model of evaluating hands, we wish to examine a few existing methods of evaluating hands. The four methods are high card points, distributional points, controls and losers. They are among the most popular methods of evaluating a hand that have been used by professional players. But through the below analysis we can see that these are not ideal ways of telling how good your hand is.

### -High Card Points

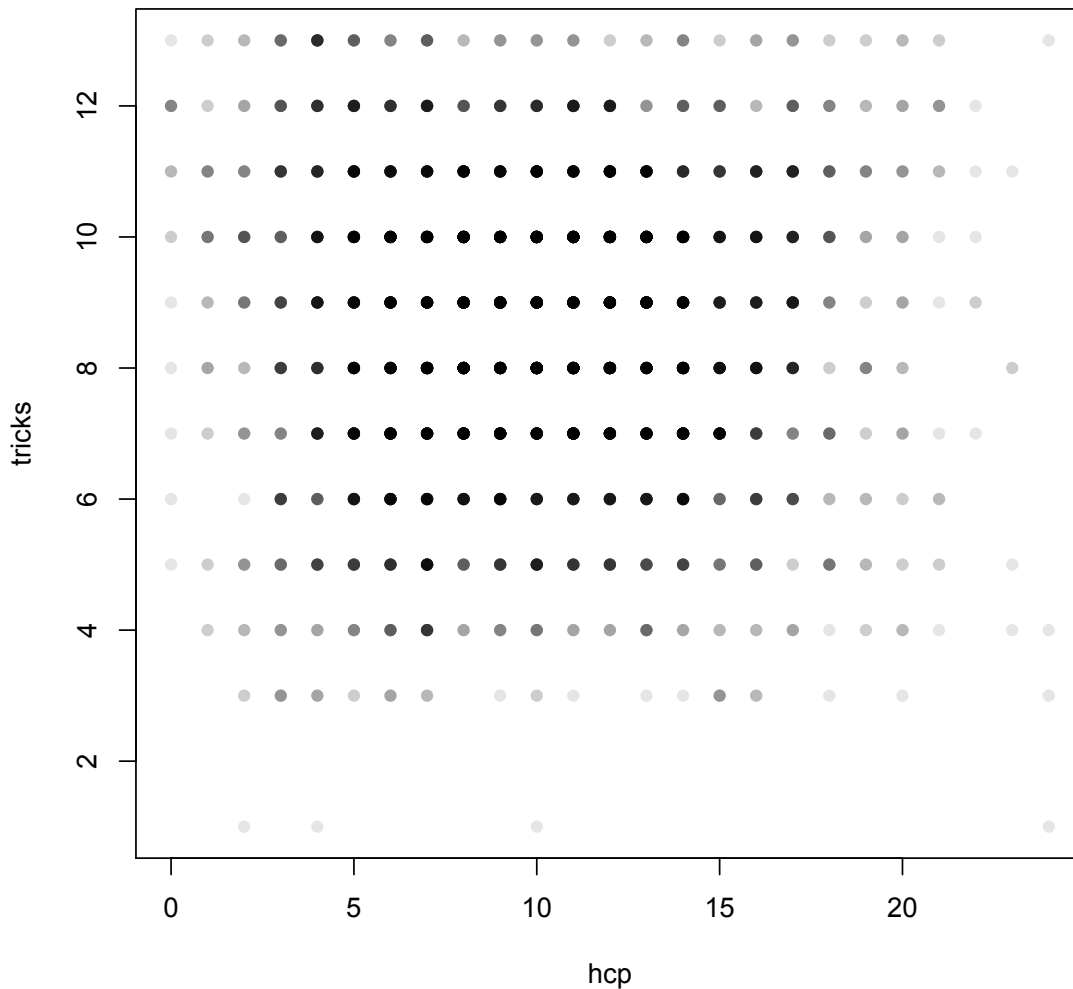
It is a classic and the most popular method of evaluating one’s hand that everyone will learn at the beginning of playing bridge. Ace is counted as 4 points, King 3, Queen 2 and Jack 1. In total there are 40 points and a hand of more than 10 points is considered good.

Here is a histogram of high card points of 1000 random sampled deals. We can see that the distribution is roughly normal and thus a robust method. However, both the plot and correlation between HCP and tricks imply that this is not so precise. And this is exactly what many bridge experts have argued: HCP does not take hand shape into account; Queens and Jacks are overestimated since they are not so powerful in the real play since they are likely to be covered by Aces and Kings.

**Histogram for hcp**



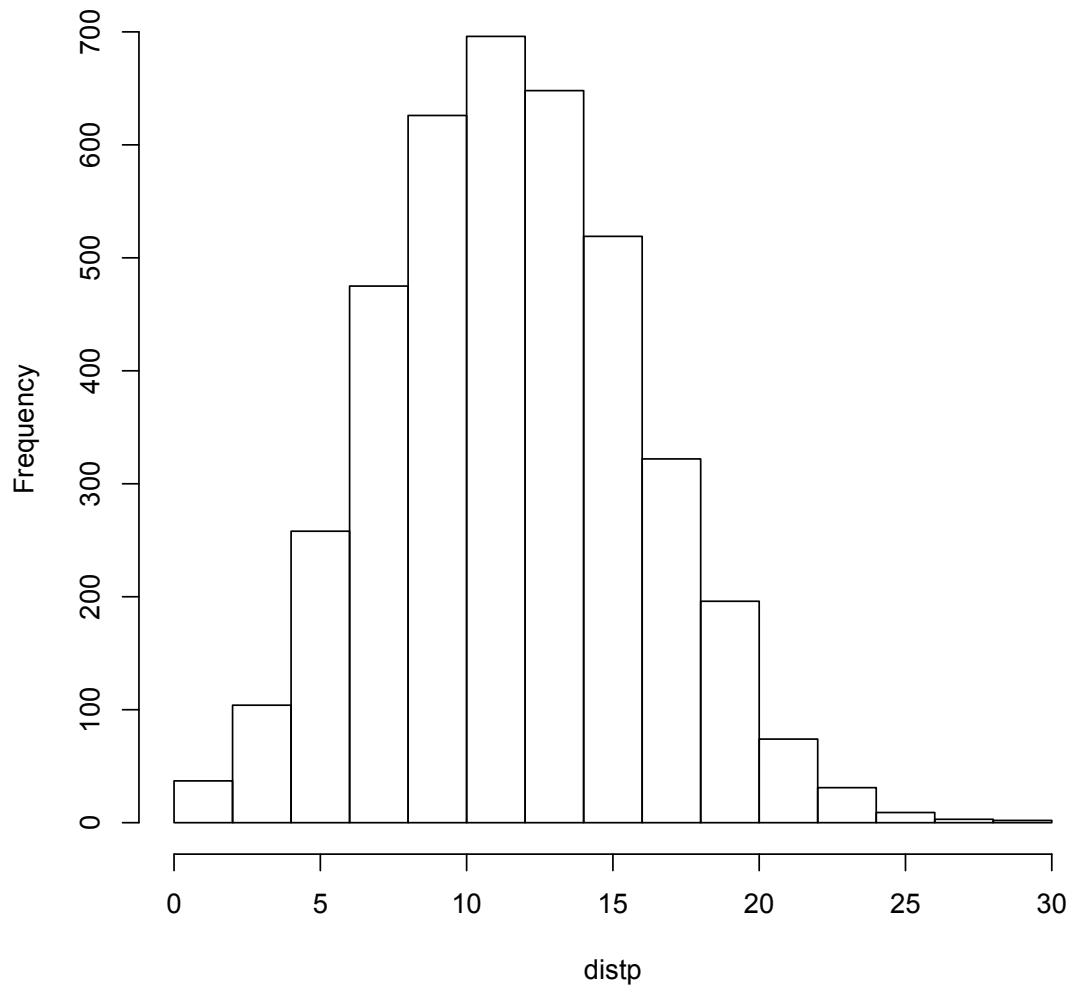
hcp vs tricks



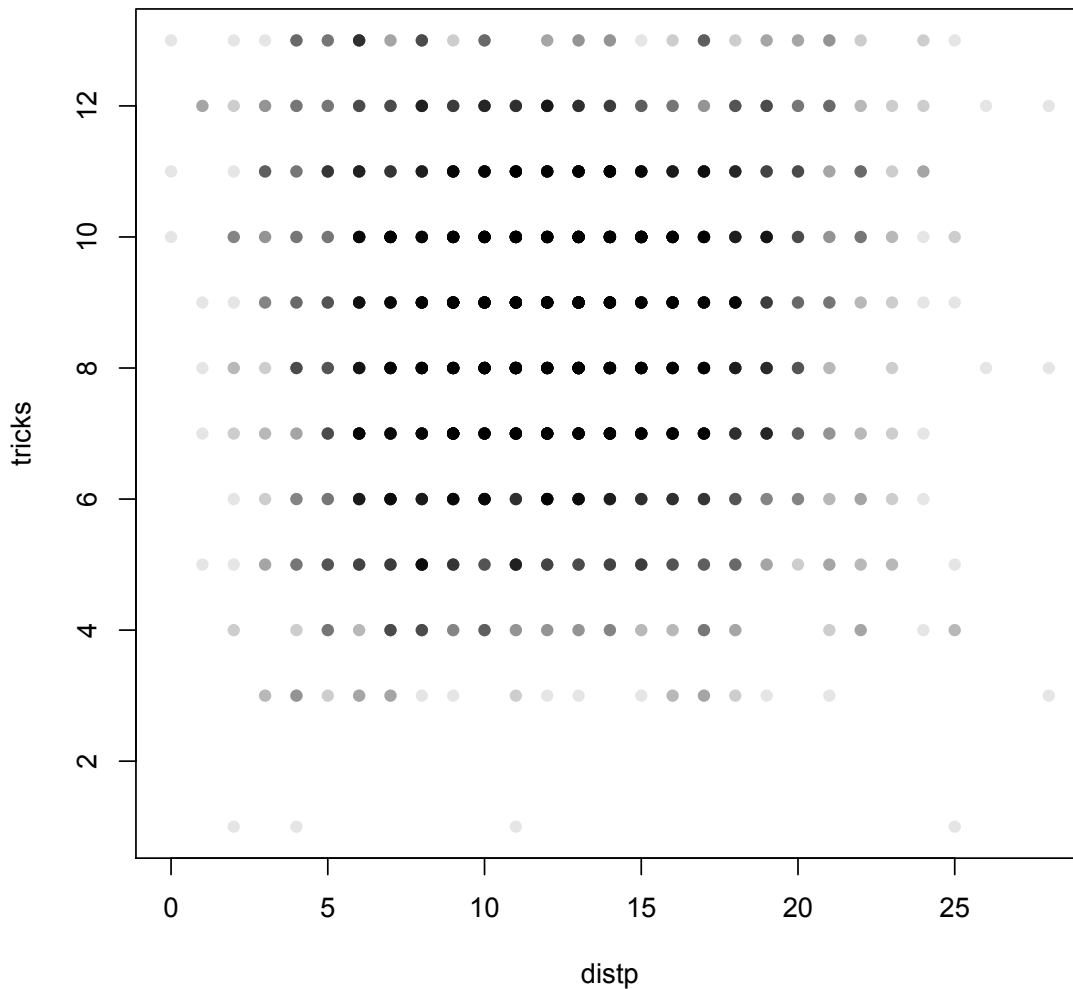
**- Distribution points**

A revised version of HCP is distribution point: that is to add distributional values to original high card points of a hand. A doubleton is worth 1 point more, singleton 3 and void 5. Since it basically inherits from HCP, we can see that the distribution is still normal-like. The correlation of distribution point and number of tricks to win is around 0.29, which is larger than before. Therefore it is better than HCP in some sense while still not good enough.

**Histogram for distp**



distp vs tricks

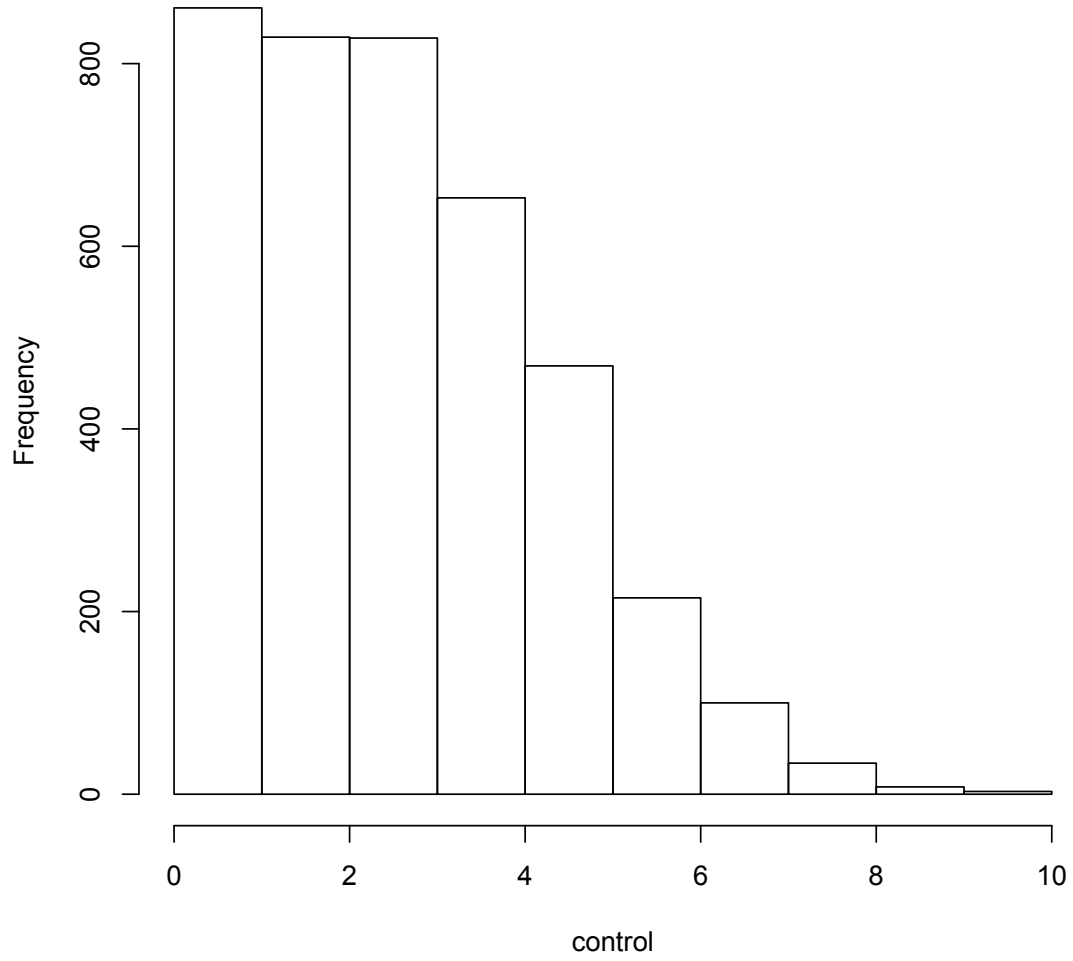


### Controls

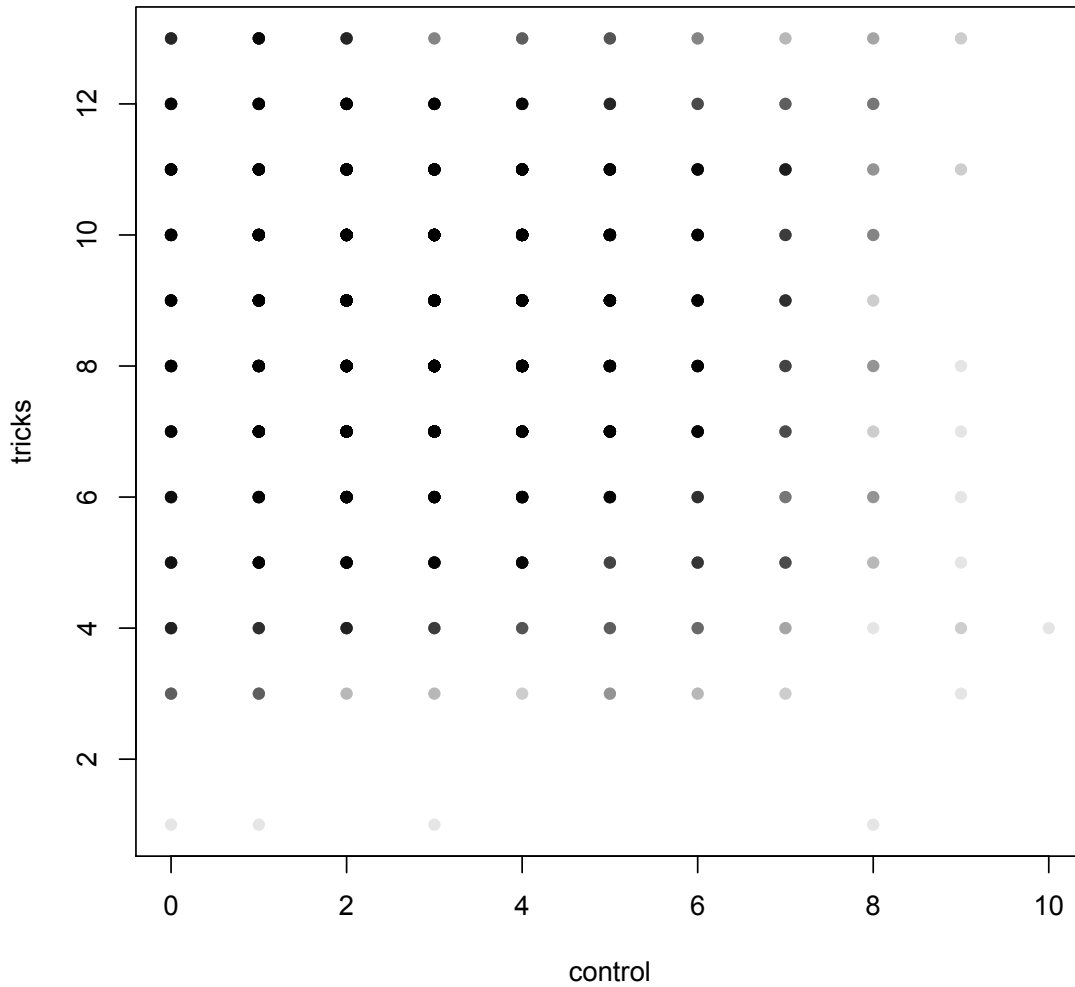
Since there are controversial ideas on values of Queen and Jacks, experts reevaluate their hands using controls when they are thinking about slam contracts. Ace is counted as two controls and King as one. But from the histogram we can see that most hands have 0-2 controls and are considered a bad hand. Because it is more like a baby version of HCP not considering Queens and Jacks, the correlation between controls and tricks won are around 0.08 which implies it is not a good indicator of

trick winning power.

**Histogram for control**



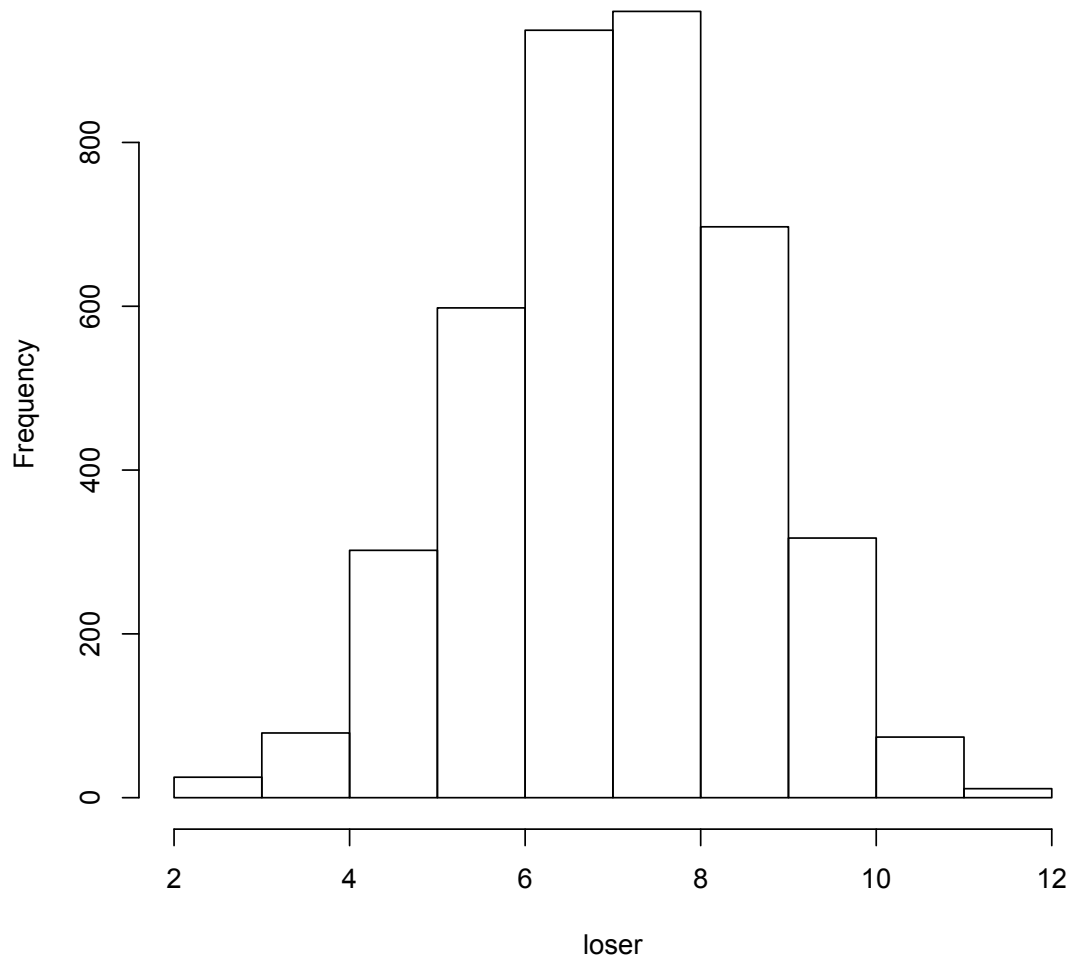
controls vs tricks



### - Losers

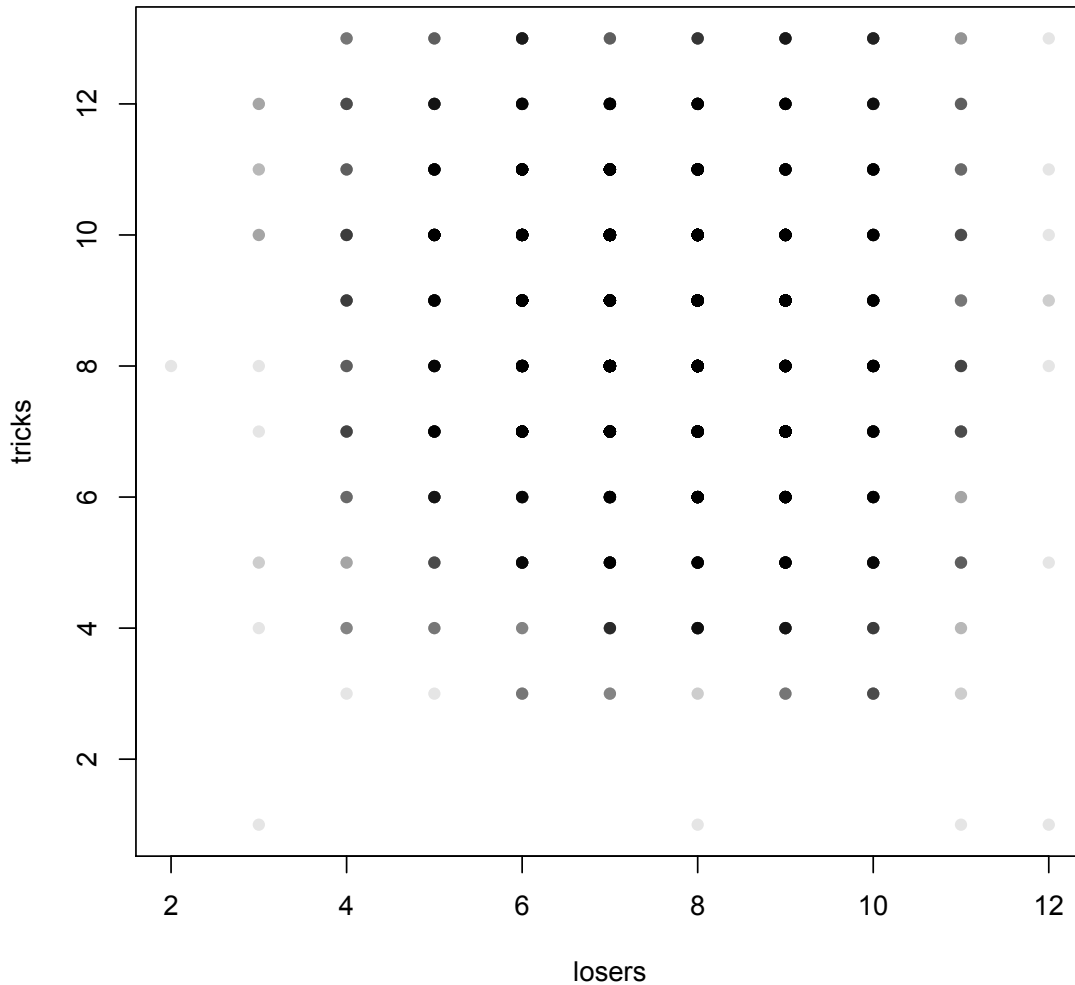
This is the last method we are going to look at. Losers are somehow more complicated to count. For each suit, just look at the three largest cards in the suit are count how many of Ace King and Queen is missing. For example, AQ has one loser, KQ432 has one loser, and Q82 has two losers. The more losers one have, the worse his hand is. The correlation is around -0.2, a decent one, but still not good enough.

**Histogram for losers**





losers vs tricks



### 3. Hand Evaluator Model

After showing that existing methods are not ideal weapons to use, we now want to come up with something more. First we want to decide which factors we are looking at in the model. Generally we have four aspects that we are concerning about a hand when we are bidding: no trump contracts and trump contracts, and defending and declaring in each case. In the program, we are going to use abbreviation, "NTOffense", "NTDefense", "TrumpOffense" and "TrumpDefense" to represent these four aspects.

We assume that the value of a hand is the linear sum of the values of each suit, and thus we have the following equation, where  $t$  is the number of tricks you expect to get in a certain suit.

$$V(h) = \sum_{s \in h} t(s)$$

We all know that a longer suit is more likely to win more tricks since we have potential chance to run the suit and win the rest of the small cards in the suit. Thus we wish to normalize and reduce the effect of this suit. We wish to take the expected winning tricks of the suit in this pattern (that is, suits that have the same high cards), minus the expected winning tricks among all the suits of the same length. The equation is stated below:

$$t(s) = ave(s) - ave(length(s))$$

Based on these two equations, we are able to analyze the expected tricks we can win given a hand. We wrote the code in python and with an input, the hand analysis function would print out expected tricks to win in four occasions respectively.



For example, I just randomly picked a hand from my hand history and wish to know how good it is. Thus I input the hand in terminal and run the program.

```
>>> hand = ["64", "AQ62", "Q762", "AT3"]
>>> hand_analysis(hand)
Expected Tricks to win:  [NTOffense, NTDefense, TrumpOffense, TrumpDefense]
shape 2-4-4-3 [6.06297777397487, 8.06529288402332, 6.895159740978901, 8.380264090447508]
Spade: 64 [-0.49743003512702266, -0.29323212061875736, -0.43250087642580226, -0.28303526024480696]
Heart: AQ62 [0.9955599318720632, 0.7781444305684797, 0.9907959606507912, 0.7607293874209571]
Diamond: Q762 [-0.45126805376184276, -0.3748388875093367, -0.47036100304572415, -0.3780265233662359]
Club: AT3 [0.8666728691869485, 0.6104258282007464, 0.8098583660238532, 0.582496953276685]
Total: [6.976512486145016, 8.785792134664451, 7.792952188182019, 9.062428647534107]
```

We can see from the output that we are expected to win 7 tricks in NT contracts, and 7.8 tricks in trump contracts if we are declarer. And we are expected to win 8.8 tricks in NT contracts and 9 tricks in trump contracts if we are defending. This is suggesting that our hand is better than average, so we should definitely open the bidding. However, it is a better hand in defense, and has a lot of potential in defending opponents' contract. So we are also happy to defend if the opponents start to interfere in the bidding.

#### **4. Strength and Weakness of the Model:**

This model uses the whole set of data out of over 700000 hands, and thus the numerical analysis is indeed reliable. From the above analysis we are able to make decisions in bidding more confidently. However, we have not considered partner's hand in the model, which means that we eliminate half of the information. Furthermore, during the bidding, as we know more and more about partner's hand, we should be able to dynamically condition on the known fact to reevaluate our hand's value, but unfortunately we have not found good ways to get that involved.

Finally, in the model, we assume the linearity of honors in different suits so that we can sum them together to get the final result. However, in the real world, a trump suit will play a larger role than other suit, but we cannot decide the trump suit just looking at our hand since partner's hand is unclear. Also we just ignore the complex correlation between honors in a suit and thus it makes the model inaccurate. In conclusion, I will recommend this model more as a reference for reconsidering your hand value than an accurate trick winning power report of a single hand.

#### **5. Reference**

Matt Ginsberg's paper on Law of total tricks:

<http://bocosan.tripod.com/ginsberg/total.HTML>

Explanations of binary file of double dummy library:

[http://bocosan.tripod.com/ginsberg/library\\_notes.HTML](http://bocosan.tripod.com/ginsberg/library_notes.HTML)